AMHERST SYSTEMS INC.

ATR SENSOR DEVELOPMENT

TECHNICAL REPORT

January 31, 1995 - February 28, 1995

Document Control No.: 614-9160002

Contract No.: N00014-95-C-0030

CLIN: 0001AA

28 February 1995

Prepared for:

Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217-5660

Prepared by:

Amherst Systems Inc.
Buffalo, NY 14221-7026
(716) 631-0610

AMHERST SYSTEMS INC.

ATR SENSOR DEVELOPMENT

TECHNICAL REPORT

January 31, 1995 - February 28, 1995

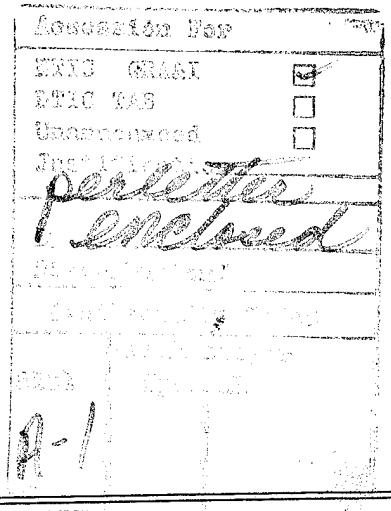
Prepared for: Office of Naval Research

Contract No.: N00014-95-C-0030

CLIN: 0001AA

Sugata Ghosal
Senior Research Engineer

Daniel G. Sentz
Manager Quality Assurance



The Contractor, Amherst Systems Inc., hereby certifies that, to the best of its knowledge and belief, the technical data delivered herewith under Contract No. N00014-95-C-0030, is complete, accurate, and complies with all requirements of the contract.

3/9/95
Date

Sugata Ghosal
Sugata Ghosal
Senior Research Engineer

Section 1
PHASE I OUTLINE

1.1 Objectives

The objectives of this Phase I effort are (1) to develop a detailed design of a smart focal plane array (FPA) for Hierarchical Foveal Machine Vision (HFMV) systems, and (2) to investigate processor node and communications interconnection design issues for medium and fine grain HFMV multiprocessors (i.e., foveal polygons). While custom VLSI is being considered for the foveal FPA, low-cost off-the-shelf components are being considered for the initial polygon architecture. The phase I work plan consists of the following tasks:

1. Multi-resolution analysis of ATR scenarios.
2. Smart foveal FPA design.
3. Foveal polygon multiprocessor design.
4. Smart foveal sensor top level design.

Section 2
WORK PERFORMED DURING PERFORMANCE PERIOD:
JANUARY 31, 1995 - FEBRUARY 28, 1995

2.1 Multi-resolution Analysis of ATR Scenarios

Multi-resolution analysis of typical ATR scenes has been performed based on the human detection and recognition performance of tactical targets presented on monochrome imaging equipment. Initial estimates of the acuity profile of foveal sensors for ATR applications can be specified according to these human visual models.

Based on human visual acuity studies by the National Transportation Safety board, an HFMV system with biomimetic detection and recognition requires four rings about the fovea, for a total of four octaves in acuity and a bandwidth compression factor of 64. Wider FOV and higher rexel count are obtained by increasing the lattice subdivision factor; doing so does not change the number of rings nor the bandwidth compression factor. The number of receptive fields in this retinotopology is 1.56% that of a uniform acuity array with the same FOV and recognition resolution throughout the FOV, giving a 98.44% reduction in retinotopology complexity.

According to the Johnson's criteria, the largest ratio of detection (peripheral) acuity to identification (foveal) acuity is roughly 1:8. Thus, an HFMV design based on the Johnson's criteria consists of three rings about the fovea for a total of three octaves in acuity and a bandwidth compression factor of 19.7. The number of receptive fields according to this design rule is 5% that of a uniform acuity array with the same FOV and recognition resolution throughout the FOV, giving a 95% reduction in retinotopology complexity. MARSAM, Bailey-Rand, and Raches models have also been investigated and they confirm the Johnson's criteria for different ATR scenario conditions. Color and motion exploitation are expected to permit lower peripheral acuity and greater bandwidth compression.

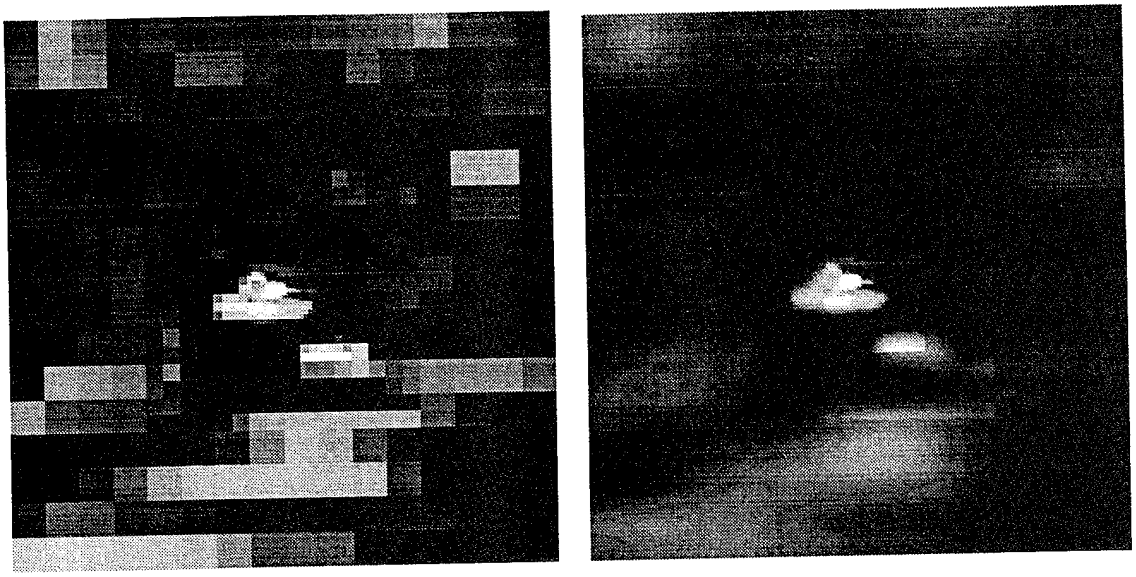
2.2. Smart Foveal FPA Design

While acuity profile specifies the retinotopology of a foveal sensor, foveal receptive fields control the retention of visual information at reduced acuity. We are investigating two classes of receptive fields: (1) *Uniform non-overlapping*, and (2) *Non-uniform overlapping*. Uniform disjoint averaging of sensor data is simpler to implement in VLSI and yields quad-tree multi-resolution representation, but the lack of convergence and divergence in peripheral receptive fields leads to aliasing problems. Non-uniform overlapping receptive fields, on the other hand, generate fewer problems with aliasing but complicate the VLSI design. Three types of overlapping receptive fields are being considered: (1) *Gaussian*, (2) *Spline* and (*3) Smoothed aggregation-based multigrid (MG) interpolation operators.* Figures 2.2-1 and 2.2-2 show the back projected rexelized images obtained by uniform, Gaussian and spline-based receptive fields for different acuity profiles. Spline-based receptive fields obviously perform the best with respect to retaining sensor plane information in the rexelized image, but are more difficult to implement in VLSI than Gaussians and have larger spatial extents. Both Gaussian and Spline-based receptive fields can be realized using cascaded digital filters.

(a) Original 256 × 256 image *(b) Uniform receptive field*

(c) Gaussian receptive field *(d) Spline-based receptive field*

Figure 2.2-1 Effect of receptive fields on rexelization. Rexelized images are back projected onto the sensor plane and displayed for three acuity rings about the fovea (four acuity levels)



(a) Uniform receptive field *(b) Gaussian receptive field*

(c) Spline-based receptive field

Figure 2.2-2 Rexelized images for four acuity rings about the fovea

Smoothed aggregation-based MG transfer operators are almost piecewise linear in nature with small spatial extent contrary to oscillating spline operators. Such transfer operators have been found very effective for iteratively solving partial differential equations. Efficiency of iterative solvers depends on the ability of transfer operators to approximate error vectors in different resolutions.

2.3. Polygon Multiprocessor Design

Several processors and multiprocessor configurations are being evaluated for use as the foveal polygon processing element (PE). The ideal PE for implementing the polygon topology

interfaces with nine nodes: one parent, four siblings, and four children nodes. However, because there is no off-the-shelf processor with this many interfaces, the polygon topology can only be approximated (under the off-the-shelf constraint). Additionally, the polygon must be scalable and must have sufficient computing power to process multi-acuity images at real-time data rates.

A PE based on the Texas Instruments TMS320C40 (C40) multidimensional parallel DSP has been selected to serve as the building block for polygon multiprocessors. The C40 is specifically designed for fine grain multiprocessors such as hypercubes, and has six high speed interfaces (each byte-parallel 20 Megabytes/sec with independent on-chip DMA) which permit a C40 to communicate simultaneously with six other C40 nodes while performing local computations. The first version of the C40 DSP was introduced in 1991. Since then, its clock speed has increased, its price has dropped, and a wide array of fine grain software development tools (simulators, compilers, debuggers, DSP libraries) has become available, making this chip an established low risk component.

Other building blocks which have been considered include the INMOS/SGS-Thompson's T9000 Transputer, the Intel I860, and the PYR-1 chip from the David Sarnoff Laboratories' spin-off Sensor. The Transputer has four interfaces which are more efficient than those of the C40, but it has very weak processor performance and its availability has been delayed for several years. The PYR-1 is a pipelined pyramid data structure generator which operates at video rates, but does not support bottom-up relinking or top-down processes. The I860 has no on-chip interfaces, and board-level inter-node communications (e.g., Quickring and crossbar switches) are slow or expensive, and still do not provide the necessary number of interconnections. A powerful alternative is a daughterboard module with a Power-PC RISC chip for processing and a transputer for communications. It is available in Europe but will not be available in the U.S. for at least one year.

The realization of the foveal polygon architecture for HFMV by a network of TMS 320C40s is currently being investigated. The design is based mainly on three considerations: regularity, feasibility, and generality. The network is designed as regularly as possible both physically and logically to facilitate the simplicity of the quasi-SIMD (single instruction multiple data) nature of the polygon structure for easier parallel programming and debugging. Ideally, for a quasi-SIMD image processing engine there should be one simple processor for each rexel. However, because general purpose microprocessors for parallel processing are generally strong in computing but relatively weak in communication, such a configuration may not produce the expected performance.

A topology which reconciles these differences is currently being investigated. The topology is based on the ratio of computing/communication power of the processors used in the system. Consider a network of processors connected as eight-connected neighbors. Assume the size of the image needed to be processed by a node is $n \times n$ and we want to convolve with kernels of size $(2k+1)$ for integer values of k. We have two choices:

1. Transfer a $(n+2k) \times (n+2k)$ subimage to each node so that no communication is needed between the nodes to achieve the convolution;

2. Transfer a n¥ n subimage to each node and use communication between the nodes to get the data needed to achieve the convolution.

For the first method, the information redundancy for each node is:

$$\alpha = \frac{4k(n+k)}{n^2}$$

at the image distribution stage. For the second method, apart from the computation, which is the same for both methods, the real-time communication that node needs to handle:

input: $4 \times (n \times k) + 4 \times (n \times k)$;

output: $4 \times (n \times k) + 4 \times (n \times k)$;

Assuming that input and output take the same amount of time per node and there is no overhead for communication management to achieve arbitrary point-to-point communication, the second method has a redundancy of

$$\alpha = \frac{8k(n+k)}{n^2}$$

It means that more communication bandwidth is needed for the second method to achieve the convolution. Particularly, if the size of the subimage n is small, the information redundancy may be very large. This illustrates that the smaller the size of the subimage on one node the more communication-intensive (rather than computation-intensive) it is to achieve some tasks. However, the communication is distributed rather than concentrated at the image transfer stage. This may reduce the image transfer bottleneck from image capture to the image processing engine.

Because we can not achieve arbitrary topology for processor networks, if the communication overhead for shorter messages and the communication management to achieve eight connections for each node is also considered, the second method is communicationally more expensive.

We can draw two conclusions based on this:

1. For a regularly connected network, increasing the number of processors in a network (smaller processing window on each node) does not automatically give good image processing performance. The performance of image processing depends on the ratio of the computation needed, the overlapping for the task and the computation/communication power ratio of the processors.

2. It is generally better off to transfer the overlapped subimages rather than to use real-time communication to achieve overlapping.

The ideal (logical) polygon is a three dimensional topology (shown in Figure 2.3-1 for the simple case of two unsubdivided rings about the fovea) representing scale-space, i.e., two dimensions are spatial (x and y of the image), and the third dimension z is acuity. The granularity of each dimension in the polygon implementation can be independently selected. This gives us considerable flexibility over which to optimize the implementation.

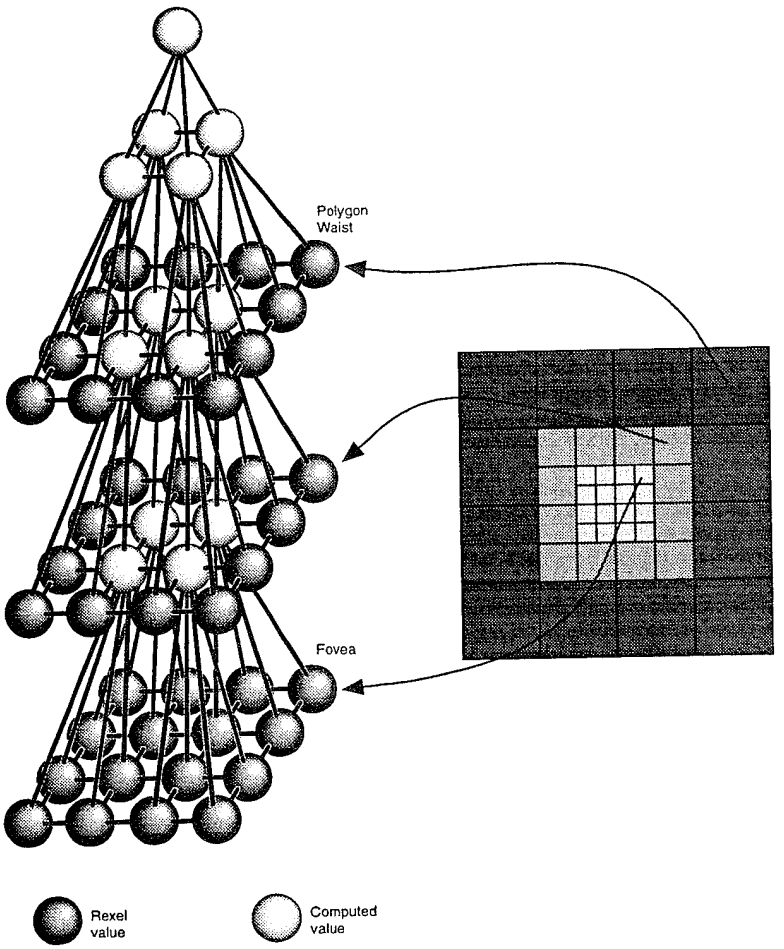


Figure 2.3-1 The foveal polygon data structure and processing hierarchy

2.4 HFMV Sensor VLSI Design

The main advantage of a foveal FPA with non-uniform sensor elements over a FPA with rexelization circuitry is that the process of generating rexel values is performed at the photo sensor level without any processing, and the complexity of the read-out mechanism (e.g., the number of lines) is reduced by the bandwidth compression factor.

The only existing foveal focal plane array is the log-polar CCD chip of Sandini and Van der Spiegel with 1,920 space variant sensing elements (30 concentric rings cut by 64 radial lines), and a small 10×10 uniform acuity fovea (Figure 2.4-1). This chip is an example of the first approach to foveal FPA design. In contrast, bi-acuity arrays consisting of a region of pixels in a wider region of 2×2 "superpixels" are examples of the second approach, because they use monolithic rexelization to build the superpixels. However, bi-acuity arrays do not support foveal machine vision in the traditional sense; true multi-resolution vision requires more than an octave in acuity, and that all acuities be treated in an SIMD fashion.

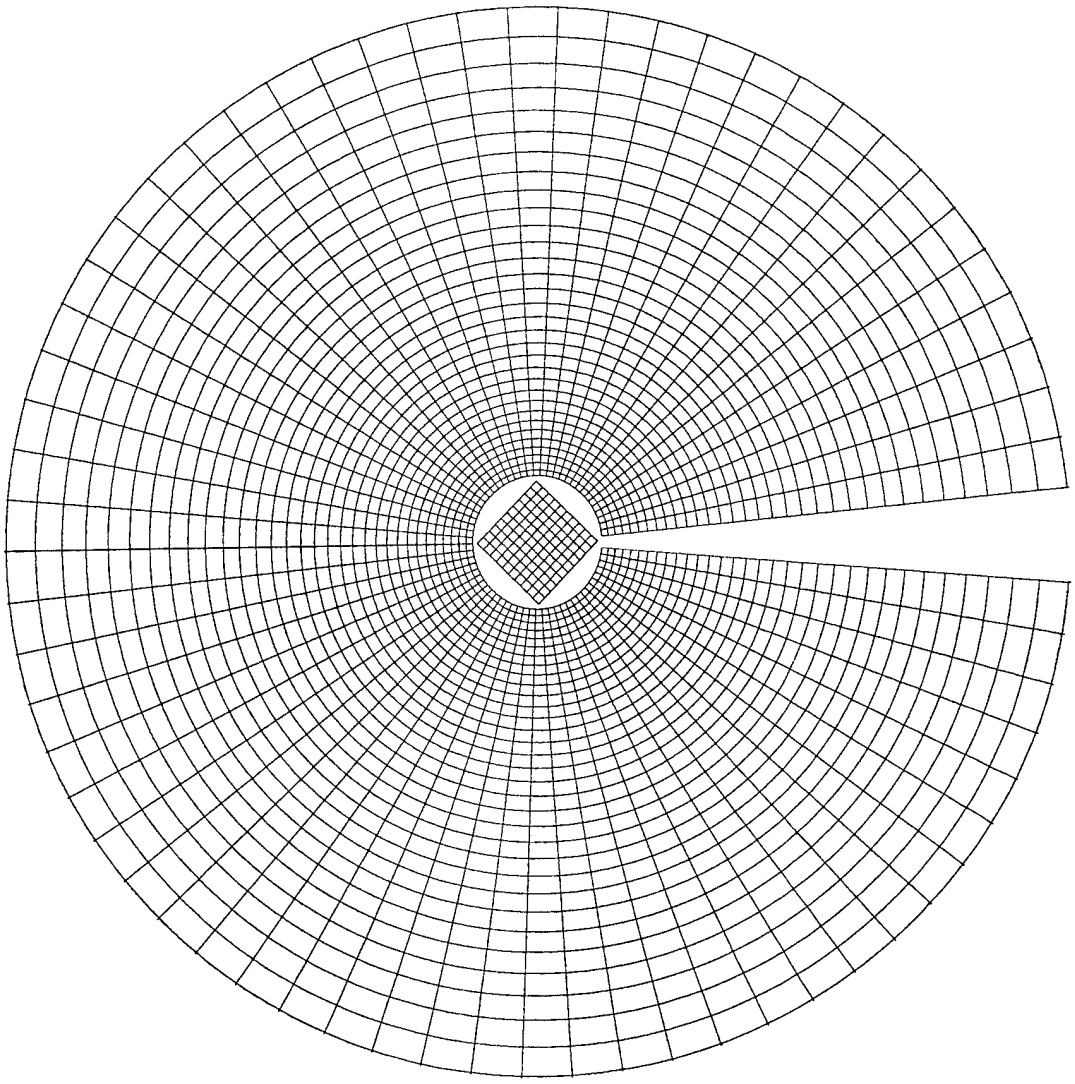


Figure 2.4-1 Sandini-Van der Spiegel log-polar foveal CCD chip

One of the most significant problems encountered by Sandini and Van der Spiegel is that the readout device of their foveal CCD chip had to accommodate a very wide dynamic range of analog signals from the sensor elements, because charge is proportional to the element's area. Peripheral receptive fields, which integrate over a wide focal plane area, will output large signals from the fovea that can be lost in the noise of a higher power readout mechanism.

Another problem encountered by Sandini and Van der Spiegel was making room for the read-out line. Instead of a blind spot near the fovea, as with human vision, their chip has a 10° radial blind sector extending from the fovea to the periphery, where read-out lines converge and continue out of the photosensitive area of the chip. Additional blind spots exist between the Cartesian fovea and the polar perifovea.

Polar sampling mimics the non-uniformity of the ganglion receptive field in the vertebrate retina. However, modeling too closely the neurophysiology which inspired foveal machine vision has precluded the commercialization of this technology, which still resides in academia.

Specifically, polar topology is very difficult to implement with current and near term technology, for the following reasons:

1. Polar topology does not follow the "Manhattan Rule" of VLSI design, which requires city block (Cartesian) circuit layout as opposed to multi-angled radial circuits.

2. Virtually all machine vision technology (e.g., imagers, algorithms, processing and data storage hardware, and development tools) features shift invariant Cartesian topology, while very few resources exist which support polar topology.

3. Hierarchical processing, prerequisite for real time active vision performance, is well understood with Cartesian topology (e.g., image pyramids and quad tree techniques), whereas no implementations exist with polar topology.

These reasons favor the HFMV topology. Nevertheless, the problems encountered by Sandini and Van der Spiegel in the development of their FPA must also be addressed in the HFMV topology. The following subsections discuss designs which solve these problems for subdivided exponential lattice foveal FPAs with variably sized sensor elements.

2.4.1 Integrated Rexel Normalization

Sandini solved the read-out mechanism dynamic range problem by using three separate read-out mechanisms, each designed to convey a limited range of charges. However, this solution increases the number of control lines and complicates the chip driver and frame grabber electronics, particularly their calibration. In fact, problems with these electronics and the calibration of some two dozen control and supply lines are preventing any further work with the chip in the robotics labs of Genoa and Pisa.

An alternative to complex readout mechanisms is to monolithically normalize the signal from a rexel by the rexel's area. Thus, all rexels in a uniformly illuminated foveal sensor will have the same value prior to read-out and digitization. However, signal normalization circuitry at the sensor elements, such as a space variant gain grid, increases chip complexity and inter-rexel spacing. Furthermore, the normalization circuitry must be very precise because the gains must be very accurate, and this requires expensive fabrication techniques, such as resistor trimming, common with precision D/A and A/D converters.

Inter-rexel spacing refers to non-transducing area in the sensor array, which is a source of sensor inefficiency (wasted flux) and signal aliasing. For these same reasons, a foveal FPA cannot be built by variably spacing uniformly sized pixels, such as placing one pixel in the center of each square of the exponential lattice, and assume its value as that of the entire rexel. This resulting decimation introduces tremendous aliasing at the perifovea and periphery; the averaging process over a rexel's receptive field is an important anti-aliasing operation.

Traditional CCD sensors transfer the entire charge to the read-out circuitry, and it is this practice that leads to the dynamic range problem in foveal CCD FPA designs. However, when phototransducing, a CCD cell is an excellent conductor, and its charge is uniformly distributed over the cell (rexel) area. It is thus possible to transfer out of the cell not the entire charge but

that of a subregion of the cell. The output signal is then the value of the rexel scaled by the ratio of the rexel area to the subregion area.

Consider three rexels, each scaled to 2^0, 2^1, and 2^2 (Figure 2.4.1-1). All of the charge of the small rexel is carried out in the traditional fashion. However, only one fourth of the charge of the next larger rexel is carried out. Likewise, only one sixteenth of the larger rexel is carried out. The signals from all three rexels, if illuminated equally, are the same. The fractional charge is isolated from the rest of the rexel by raising a potential across a barrier (dotted line). This barrier is in fact an extension of the read-out mechanism into the rexel itself, and the signal *f* that raises and lowers the potential is just one of the read-out clock phases.

This approach to normalizing rexel values has the distinct advantage of not requiring additional active circuitry because it is integrated into the rexel itself. Also, because size tolerances of VLSI features are much smaller than resistive and capacitance tolerances, the scaling is also much more accurate than that supported by active or passive circuitry, such as resistor networks. Scaling is determined by the areas of the rexel and rexel subregion, which are precisely controlled in the manufacturing process and do not require any resistor trimming.

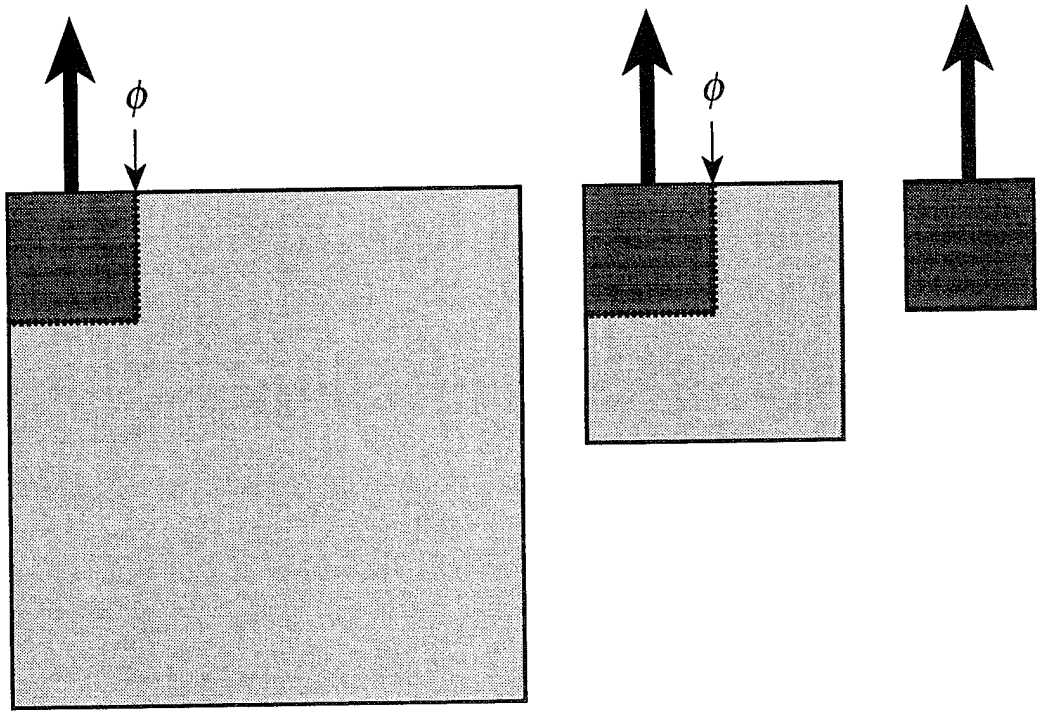


Figure 2.4.1-1 Integrated rexel value normalization

2.4.2 Foveal Read-Out Mechanisms and Blind Spots

The FPA read-out mechanism, a data communications bottleneck in most imagers, determines in great part the frame rate, the image integration time, and the spacing between sensor elements. The exponential lattice lends itself to two orthogonal foveal read-out designs, illustrated for the root exponential lattice in Figures 2.4.2-1 and 2.4.2-2. The first is a radial design which outputs an entire ring of rexels at a time. The time required to output the entire frame of rexels is proportional to the number of rings, plus time for the inner fovea rexels. This design exploits the fact that all the rexels line up radially. The second is a tangential design which outputs an entire

radial sector of rexels at a time. This design exploits the fact that all the rings have the same number of rexels, and the time required to output the entire frame of rexels is equal to the number of rexels in any given ring. The first design is favored over the second because read-out line lengths are minimized in the former, while the latter imposes a circuitous route which adds noise, signal loss, and inter-rexel spacing.

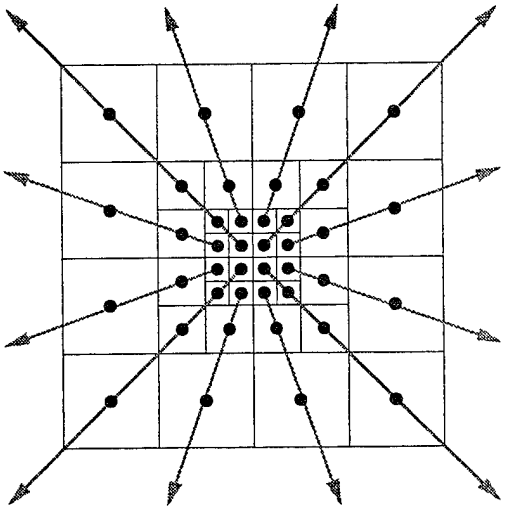


Figure 2.4.2-1 Radial foveal readout

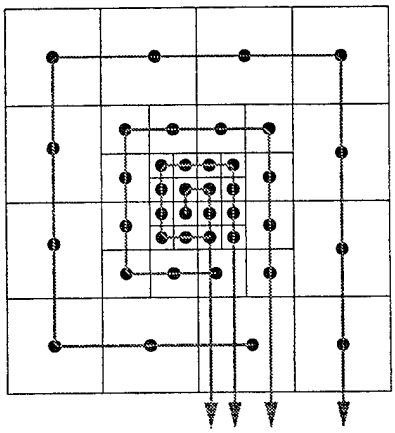


Figure 2.4.2-2 Tangential foveal readout

A read-out design is recommended which combines features of the radial topology with traditional linear parallel topologies. Specifically, read-out lines in any quadrant of the FPA will follow the Manhattan rule unidirectionally to a charge priming transfer (CPT) array at the perimeter of the photosensitive area of the chip, where buffering, A/D conversion, and possibly monolithic foveal image processing are performed. Figure 2.4.2-3 illustrates this design with a read-out line from every rexel to the CPT array. Rexels may be multiplexed radially such that one read-out line conveys the values of more than one rexel in a radial slice, as implied by Figure 2.4.2-1, to reduce inter-rexel spacing.

Multiple outputs are becoming prevalent in high pixel rate cameras. These interfaces split the overall pixel array in half or in quadrants, and each of two or four outputs conveys the imagery of one section. The inherent symmetry of the exponential lattice and the parallel-radial foveal read-out design supports two, four, and even eight outputs. In each case, the read-out circuitry of any foveal sensor output is a mirror image of that of the other outputs, so all outputs can share one set of timing and control signals, unlike the three read-out mechanisms of the aforementioned log-polar chip.

A major concern in the design of a conventional FPA is the dead-space between pixels. The interpixel spacing is always maintained as uniform and small as possible to minimize aliasing, and as uniform as possible to minimize geometric distortion in the image. The same concern applies to the design of a foveal FPA, except that the concept of uniform inter-element spacing must be clarified. Uniform inter-rexel spacing can be small relative to peripheral rexels but large relative to fovea rexels, producing aliasing which is more pronounced at the fovea, and which would interfere with recognition.

A better objective is to keep the inter-rexel spacing small and proportional to rexel size. This objective is met with the parallel-radial read-out design. Specifically, the number of lines between rexels is proportional to the size of the rexels. This is made possible by the inverse-linear acuity profile of the subdivided exponential lattice, which maintains the number of rexels (and inter-rexel pathways) constant for any major ring, while increasing rexel size proportionally with distance from the center. The geometric normalization of the read-out design also makes the design scalable, i.e., it can accommodate any number of rings, and any number of outer rings can be introduced without adding additional inter-rexel spacing to the interior rings.

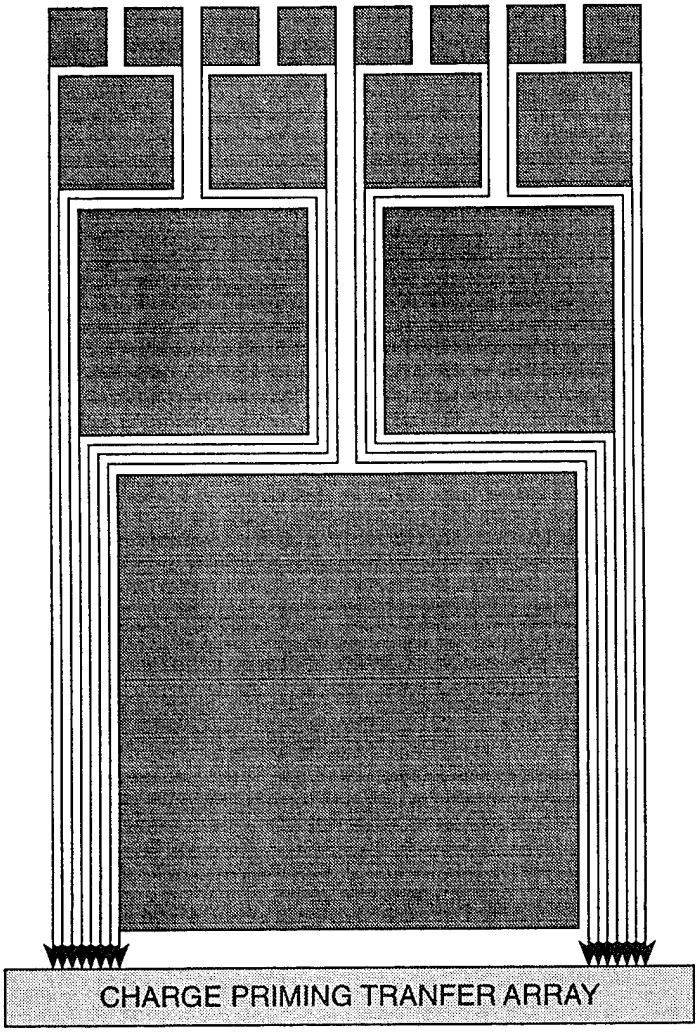


Figure 2.4.2-3 Scalable parallel-radial read-out with acuity normalized inter-rexel spacing

Section 3
PLANNED WORK FOR NEXT PERFORMANCE PERIOD

We will investigate multigrid transfer operators for the design of foveal receptive fields. VLSI implementability will be addressed for monolithic realization of different receptive field designs. Hierarchical optical flow and segmentation performance are being considered for evaluating the direct merits of different receptive fields for the processing of sensor data in the ATR domain. We will also be investigating VLSI structures for the readout mechanism, which are relatively independent of receptive field shape. Physical and logical polygon architectures which reconcile the tradeoff between computation and communication are currently being evaluated. We will examine mapping the logical polygon onto a 2-D and a 3-D architecture during next performance periods.

Section 4
AREAS OF CONCERN

At this time, there are no areas of concern. Staffing projections indicate a timely completion of this Phase I effort.

OFFICE OF THE UNDER SECRETARY OF DEFENSE (ACQUISITION)
DEFENSE TECHNICAL INFORMATION CENTER
CAMERON STATION
ALEXANDRIA, VIRGINIA 22304-6145

IN REPLY REFER TO DTIC-OCC

SUBJECT: Distribution Statements on Technical Documents

TO: OFFICE OF NAVAL RESEARCH
CORPORATE PROGRAMS DIVISION
ONR 353
800 NORTH QUINCY STREET
ARLINGTON, VA 22217-5660

1995 1027 024

1. Reference: DoD Directive 5230.24, Distribution Statements on Technical Documents, 18 Mar 87.

2. The Defense Technical Information Center received the enclosed report (referenced below) which is not marked in accordance with the above reference.

TECHNICAL REPORT
N00014-95-C-0030
TITLE: ATR SENSOR DEVELOPMENT

3. We request the appropriate distribution statement be assigned and the report returned to DTIC within 5 working days.

4. Approved distribution statements are listed on the reverse of this letter. If you have any questions regarding these statements, call DTIC's Cataloging Branch, (703) 274-6837.

FOR THE ADMINISTRATOR:

1 Encl

GOPALAKRISHNAN NAIR
Chief, Cataloging Branch

FL-171
Jul 93

DISTRIBUTION STATEMENT A:

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION IS UNLIMITED

DISTRIBUTION STATEMENT B:

DISTRIBUTION AUTHORIZED TO U.S. GOVERNMENT AGENCIES ONLY; (Indicate Reason and Date Below). OTHER REQUESTS FOR THIS DOCUMENT SHALL BE REFERRED TO (Indicate Controlling DoD Office Below).

DISTRIBUTION STATEMENT C:

DISTRIBUTION AUTHORIZED TO U.S. GOVERNMENT AGENCIES AND THEIR CONTRACTORS; (Indicate Reason and Date Below). OTHER REQUESTS FOR THIS DOCUMENT SHALL BE REFERRED TO (Indicate Controlling DoD Office Below).

DISTRIBUTION STATEMENT D:

DISTRIBUTION AUTHORIZED TO DOD AND U.S. DOD CONTRACTORS ONLY; (Indicate Reason and Date Below). OTHER REQUESTS SHALL BE REFERRED TO (Indicate Controlling DoD Office Below).

DISTRIBUTION STATEMENT E:

DISTRIBUTION AUTHORIZED TO DOD COMPONENTS ONLY; (Indicate Reason and Date Below). OTHER REQUESTS SHALL BE REFERRED TO (Indicate Controlling DoD Office Below).

DISTRIBUTION STATEMENT F:

FURTHER DISSEMINATION ONLY AS DIRECTED BY (Indicate Controlling DoD Office and Date Below) or HIGHER DOD AUTHORITY.

DISTRIBUTION STATEMENT X:

DISTRIBUTION AUTHORIZED TO U.S. GOVERNMENT AGENCIES AND PRIVATE INDIVIDUALS OR ENTERPRISES ELIGIBLE TO OBTAIN EXPORT-CONTROLLED TECHNICAL DATA IN ACCORDANCE WITH DOD DIRECTIVE 5230.25, WITHHOLDING OF UNCLASSIFIED TECHNICAL DATA FROM PUBLIC DISCLOSURE, 6 Nov 1984 (Indicate date of determination). CONTROLLING DOD OFFICE IS (Indicate Controlling DoD Office).

The cited documents has been reviewed by competent authority and the following distribution statement is hereby authorized.

A

(Statement)

OFFICE OF NAVAL RESEARCH
CORPORATE PROGRAMS DIVISION
ONR 353
800 NORTH QUINCY STREET
ARLINGTON, VA 22217-5660

(Controlling DoD Office Name)

(Reason)

(Controlling DoD Office Address, City, State, Zip)

DEBRA T. HUGHES
DEPUTY DIRECTOR
CORPORATE PROGRAMS OFFICE

19 SEP 1995

(Signature & Typed Name)

(Assigning Office)

(Date Statement Assigned)